

علم اللغة والتّقنيات المعاصرة

د.المُعْتَزُّ بالله السَّعيد

moataz@cu.edu.eg

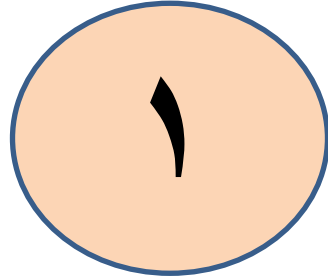
مرحبا بكم

welcome



المُحاضَرة الأولى

في حَوَسِبَةِ اللُّغَةِ



- في مفهوم حوسبة اللغة -

- إرهاصات النشأة والتطور -

- مستويات التحليل اللغوي في حوسبة اللغة -

- في مفهوم حوسبة اللغة -

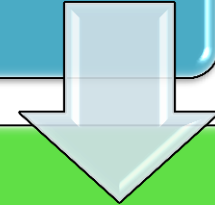
- إرهاصات النشأة والتطور -

- مستويات التحليل اللغوي في حوسبة اللغة -

في مفهوم حوسبة اللغة

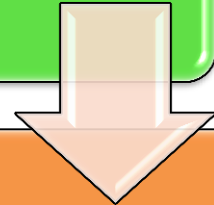
اللِّسَانِيَّات الحَاسُوبِيَّة

Computational Linguistics



مُعَالَجَة اللُّغَات الطَّبِيعِيَّة

Natural Language Processing



تقنيات اللغة الطَّبِيعِيَّة (الإنسانية)

Human Language Technology

في مفهوم حوسبة اللغة

- يُمكنُ القولُ إنّ اللّسانيّات الحاسوبية فرعٌ من فُرُوع اللّسانيّات الحديثة. غايَتُها اللُّغة ووسيلَتُها الحاسوب. وتُمثِّلُ ذلك العلم الذي يُعنى بإخضاع الآلة لمتطلّبات البحث اللُّغويّ على كافّة مُستوياته استنادًا إلى مُعطيات اللُّغات الطّبيعيّة. ويغلُبُ عليها الجانب التّنظيريّ. أمّا مُعالجة اللُّغات الطّبيعيّة، فتُعنى بتوجيه الأنظمة الحاسوبية إلى محاكاة قواعد اللُّغات الطّبيعيّة اعتمادًا على مُعطيات هذه اللُّغات، وتُمثِّلُ الجانب التّطبيقيّ من اللّسانيّات الحاسوبية. وتُشيرُ تقنياتُ اللُّغات الإنسانيّة إلى مجموعة التّطبيقات التي تُستخدَمُ في بناء وتطوير أدوات المعالجة الآليّة للُّغات الطّبيعيّة.

- في مفهوم حوسبة اللغة -

- إرهاصات النشأة والتطور -

- مستويات التحليل اللغوي في حوسبة اللغة -

إرهاصات النشأة والتطور

- ظَهَرَت اللِّسَانِيَّاتُ الحَاسُوبِيَّةُ فِي الخَمْسِينِيَّاتِ مِنَ القَرْنِ العَشْرِينَ عِبْرَ مِيَادِينَ التَّرْجَمَةِ بَيْنَ اللُّغَاتِ الْإِنْسَانِيَّةِ، عَقِبَ ظُهُورِ الْجِيلِ الْأَوَّلِ مِنَ الحَوَاسِبِ.
- أَمَكْنَ إِخْضَاعُ الْآلَةِ لِمَحَاكَاةِ مَهَارَاتِ التَّفْكِيرِ الْعُلْيَا عِنْدَ الْإِنْسَانِ [مِثْلُ: الْاسْتِنْبَاطِ، وَالتَّعَلُّمِ، وَالْحِفْظِ، وَالتَّذَكُّرِ] – الذِّكَاءِ الْإِصْطِنَاعِيِّ.
- فِي مُدَّةٍ قَصِيرَةٍ انْتَقَلَتِ تَقْنِيَّاتُ اللِّسَانِيَّاتِ الحَاسُوبِيَّةِ مِنَ الْوِلَايَاتِ الْمُتَّحِدَةِ الْأَمْرِيكِيَّةِ – مَهْدَهَا الْأَوَّلِ – إِلَى الْأَمْرِيكَتَيْنِ وَأُورُوبَا، ثُمَّ إِلَى مَنَاطِقَ أُخْرَى حَوْلَ الْعَالَمِ.

إرهاصات النشأة والتطور

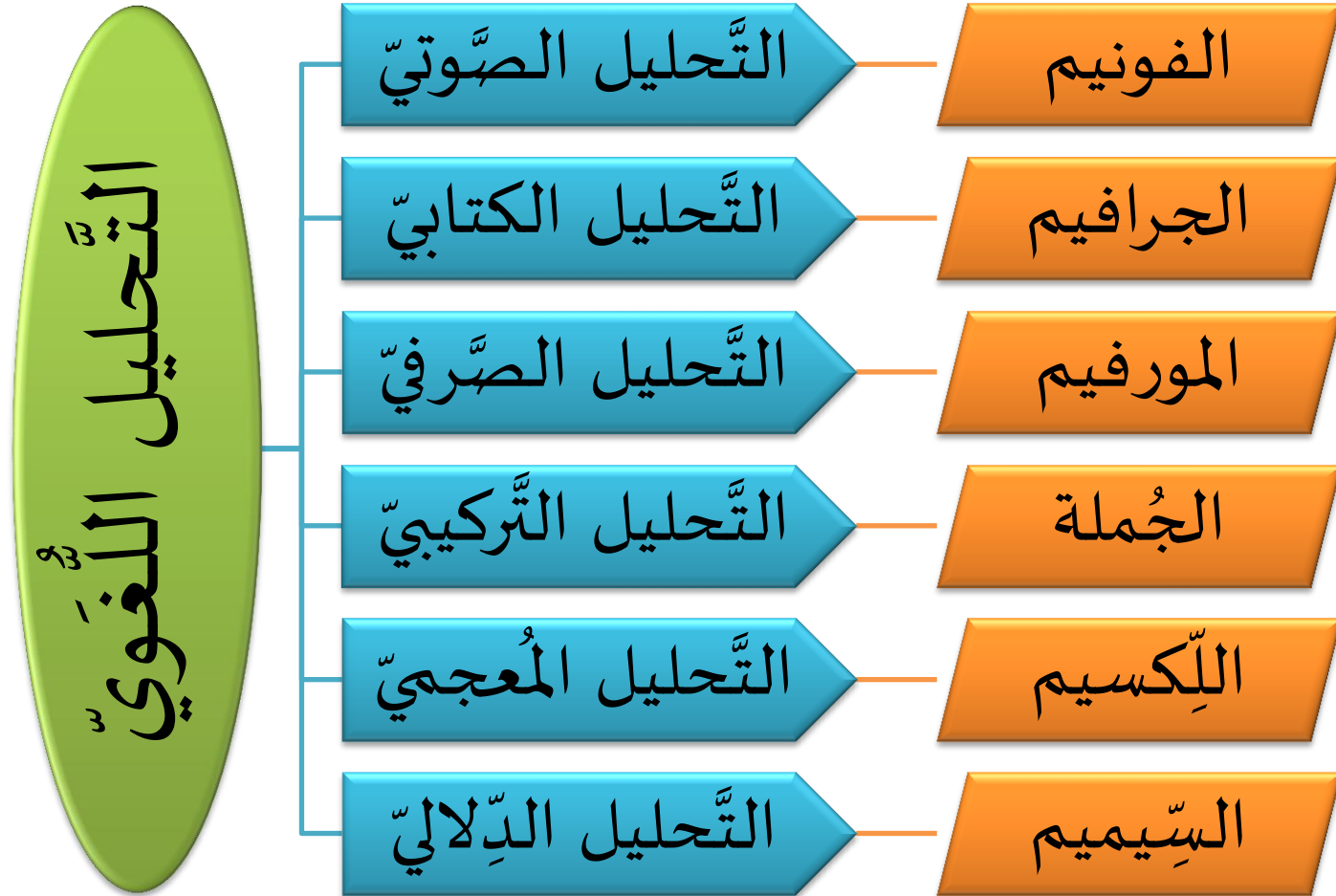
- مع التطور السريع لهذه التقنيات، ظهر ميدان جديد يُعنى بتوظيف الآلة في صناعة المعجم؛ واستقل هذا الميدان لاحقاً فيما يُعرف بـ "الصناعة المعجمية الحاسوبية".
- تُعدُّ اللغة العربية إحدى اللغات التي شملتها عناية الباحثين في حوسبة اللغة من خلال ميادين الترجمة الآلية.
- كانت الطفرة الحقيقية في الطريق إلى حوسبة اللغة العربية في مُنتصف الثمانينيات من القرن العشرين، حيثُ اتَّجه الباحثون إلى بناء وتطوير أدواتٍ لمعالجتها آلياً، كما عملت العديد من المؤسسات المعنية بحوسبة اللغة على توفير الموارد اللغوية اللازمة.

- في مفهوم حوسبة اللغة -

- إرهاصات النشأة والتطور -

- مستويات التحليل اللغوي في حوسبة اللغة -

مُستويات التَّحليل اللُّغويّ في حَوَسبة اللُّغة



مُستويات التَّحليل اللُّغويّ في حوسبة اللُّغة

• التَّحليل الصَّوتيّ:

✓ يُعنى بتحليل كلمات اللُّغة (في صُورتها المنطوقة) إلى الوحدات المُشكِّلة لأصوات الحُرُوف وما يَنْتُج عنها من تنويعاتٍ صَوْتِيَّة.

✓ وحدة التَّحليل الصَّوتيّ هي: الفونيم *Phoneme*، ويُمكن تعريفه بأنَّه: أصغر وحدة مُميِّزة للصَّوت اللُّغويّ، ولا يكونُ لهذه الوحدة معنى بمُفردها.

✓ تنقسم الفونيمات إلى: مَقْطِعيَّة، وتشتمل على كُلِّ الصَّوامت والصَّوائت في اللُّغة، وفوق مَقْطِعيَّة، وتشتمل على التَّنويعات الصَّوتِيَّة الَّتِي تظهِر في سياق الكلام، مثل: الفاصلة والنَّغمة والنَّبرة.

مُستويات التَّحليل اللُّغويِّ في حَوَسَبة اللُّغة

• التَّحليل الصَّوتيّ:

✓ تتنوعُ أشكال الفونيم إلى مجموعةٍ من الصُّور الصَّوتِيَّة (أو الأشكال النُّطقيَّة) الَّتِي تُعرَفُ بـ (ألوفونات *Allophones*)؛ ويُمكن تعريف الألفون بأنه "أحد الأشكال المُتعدِّدة للفونيم الواحد".

✓ تتَّحدُ الفونيمات مع بعضها لتُنتِجَ (المقاطع *Syllables*). وهي "وحداتٌ صوتِيَّة، تنُتِجُ عن ائتلاف الفونيمات مع بعضها".

✓ تأتي كلماتُ اللُّغة "أحاديَّة المقطع"، ومن أمثلتها في العربيَّة (إنْ، فِ)؛ أو "ثنائيَّة المقطع"، ومن أمثلتها (أبى، يَزِنُ)؛ أو "ثلاثيَّة المقطع"، ومن أمثلتها (كَتَبَ)؛ أو "مُتعدِّدة المقاطع"، ومن أمثلتها (مُسْتَوِطِنون).

مُستويات التَّحليل اللُّغويّ في حَوْسَبَة اللُّغة

• التَّحليل الكتابي:

✓ يُعنى بتحليل كلمات اللُّغة (في صُورَتِها المكتوبة) إلى الوحدات الكتابيّة المُشكَّلة للحُرُوف، سواءً أكانت مخطوطةً أم مطبوعة؛ وبعبارةٍ أخرى، يُعنى بتحليل النِّظام الكتابي للُّغة إلى وحداتٍ كتابيّة.

✓ وحدة التَّحليل الكتابي هي: الجرافيم *Grapheme*، ويُمكن تعريفه بأنّه: أصلُ وحدات اللُّغات المكتوبة.

✓ يتشكَّل الجرافيمُ الواحد من (الألوجرافات *Allographs*)، وهي مجموعةٌ من الأشكال الَّتِي يَتَّخِذُها الجرافيم [المحرّف] بحسب طريقة كتابته، أو موضعه في الكلمة.

مُستويات التَّحليل اللُّغويّ في حَوَسبة اللُّغة

• التَّحليل الكتابي:

<i>Allographs</i>				<i>Graphemes</i>			
ؤ	أ	ئ	ء	@	،	١	ء
ب	ب	ب	ب	#	؛	٢	ب
ت	ت	ت	ت	\$	؟	٣	ت
ث	ث	ث	ث	%	!	٤	ث

مُستويات التَّحليل اللُّغويِّ في حَوْسَبَةِ اللُّغة

• التَّحليل الصَّرْفِيّ:

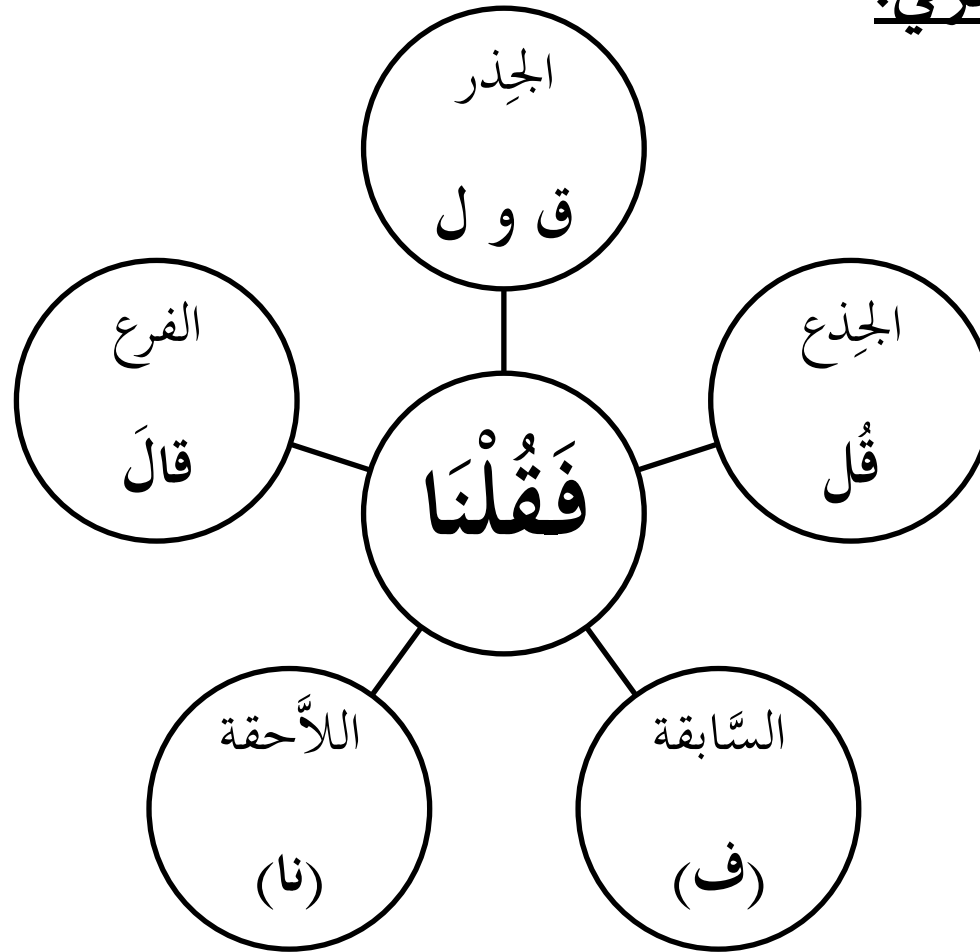
✓ يُعْنَى بِتَحْلِيلِ الْوَحْدَاتِ التَّصْرِيفِيَّةِ (الصَّرْفِيَّةِ / الْبِنَوِيَّةِ) الْمُجَرَّدَةِ الدَّالَّةِ عَلَى مَعْنَى، كَمَا يُعْنَى بِتَحْلِيلِ أَجْزَاءِ الْكَلِمَةِ النَّاتِجَةِ عَنْ تَجْمُّعِ الْمُورَفِيمَاتِ مَعَ بَعْضِهَا أَوْ تَفَكُّكِهَا.

✓ وَحْدَةُ التَّحْلِيلِ التَّصْرِيفِيِّ هِيَ الْمُورَفِيم *Morpheme* الَّذِي يُعْرَفُ - فِي اصْطِلَاحِ اللُّغَوِيِّينَ - بِأَنَّهُ أَصْغَرُ وَحْدَةٍ لُغَوِيَّةٍ مُجَرَّدَةٍ لَهَا مَعْنَى.

✓ تَتَدَاخَلُ الْمُورَفِيمَاتُ نَتِيجَةَ التَّأَلُّفِ بَيْنَ الْعُنَاصِرِ الَّتِي تَتَكَوَّنُ مِنْهَا الْكَلِمَةُ (الْجَذْرُ، وَالْجَذْعُ، وَالْفَرْعُ، وَالسَّابِقَةُ، وَاللَّاحِقَةُ).

مُستويات التَّحليل اللُّغويّ في حوسبة اللُّغة

• التَّحليل الصَّرفيّ:



مُستويات التَّحليل اللُّغويِّ في حَوَسَبة اللُّغة

• التَّحليل التَّركيبيّ:

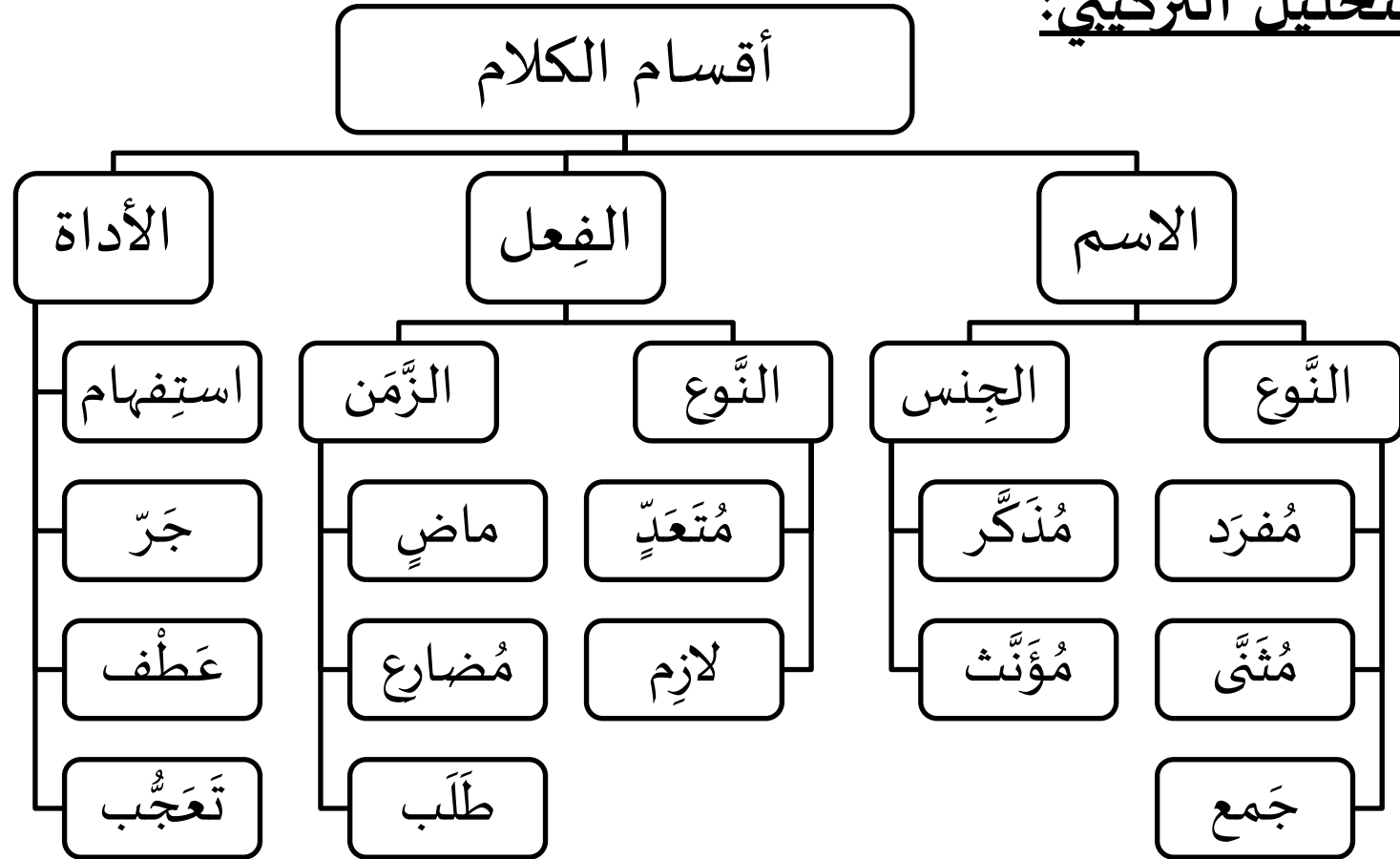
✓ يُعنى بتحليل الوحدات المُكوِّنة للتَّركيب النَّحويِّ، والتي تُعرَفُ بـ "أقسام الكلام *Parts of Speech*".

✓ يتناولُ جانبي "الوصف النَّحويِّ" الَّذي يُميِّزُ بين العلاقات الدَّاخليَّة لعناصر الكلام، و "الإعراب" الَّذي يضبطُ حركات الكلام.

✓ يختلفُ العُلَماءُ في تحديد وحدة التَّحليل التَّركيبيِّ؛ فمنهم من يراها "الفِقرة *Clause*"، ومنهم مَنْ يراها "العِبارة *Phrase*"، ومنهم مَنْ يراها "الجُملة *Sentence*"; وهذا الرَّأي الأخيرُ هو الأكثرُ شيوعًا والأقربُ إلى منطق اللُّغة.

مُستويات التَّحليل اللُّغويِّ في حَوْسَبَةِ اللُّغة

• التَّحليل التَّركيبي:



مُستويات التَّحليل اللُّغويِّ في حوسِّبة اللُّغة

• التَّحليل المُعجمي:

✓ يُعنى بتحليل الوَحَدات المعجميَّة الَّتِي تَتَشَكَّلُ عن المداخل الرِّئيسيَّة للمعجم؛ ويبحثُ في الثَّروة اللَّفظيَّة لِلُّغة وطريقة جمع المُفردات.

✓ وحدة التَّحليل المعجميَّ هي: اللِّكسيم *Lexeme*، ويُعرَّفُ بأنَّه: "الوحدة المُفتاحيَّة الَّتِي تُشكِّلُ قوائِمُها مداخلَ المعجم".

✓ تتألَّفُ مادَّةُ المعجم اللُّغويِّ من المداخل والوحدات المعجميَّة؛ حيثُ يُمثِّلُ المدخل المعجميُّ الحَقْلَ الَّذِي تنتهي إليه مجموعةُ الكلمات المُشتركة في مادَّةٍ لُغويَّةٍ واحدَةٍ، وتُمثِّلُ الوحدة مجموعةُ الكَلِمات الرِّاسيَّة الَّتِي تُشكِّلُ القوائِمَ المُنسدِلَةَ عن المدخل المعجميَّ.

مُستويات التَّحليل اللُّغويّ في حَوَسبة اللُّغة

• التَّحليل المُعجميّ:

✓ المعنى المعجميّ هُوَ وَسيلةُ شَرْحِ الوحدات المُعجميّة اعْتِمادًا على دلالاتها اللُّغويّة. ويأتي على إحدى صُورَتَيْن: المعنى الحَقِيقِيّ، وهو الَّذي تُشيرُ إليه مَدلُّولات الوحدات المُعجميّة، والمعنى المجازيّ، وهو المعنى الَّذي يَرمي إلى مَعْنَى مُخْتَلِفٍ عن المعنى الحَقِيقِيّ، وتَدُلُّ عليه قَرينةٌ لَفْظيّةٌ أو مَعنويّة.

✓ تَتَعَدَّدُ المعاني المُعجميّةُ لِلْمُفْرَداتِ بِتَعَدُّدِ التَّنَويعاتِ السِّياقيّةِ وأَساليبِ التَّعبيرِ. [كَلِمَة (عَيْن): عَضو الإِبصار، الجاسُوس، ...].

✓ يُعْنى التَّوصيفُ النَّحْويُّ لِلوَحَداتِ المُعجميّةِ بِالْجانبِ السَّماعيِّ.

مُستويات التَّحليل اللُّغويّ في حوسبة اللُّغة

• التَّحليل الدِّلاليّ:

- ✓ يُعنى بتحليل الوحدات الدِّلاليَّة وبيان دلالاتها ووظائفها في فهم المعنى، سواءً أكانَ هذا المعنى في اللُّغة المكتوبة أم اللُّغة المنطوقة.
- ✓ وحدة التَّحليل الدِّلاليّ هي: السِّيميم *Sememe*؛ ونعني بها أصغر وحدة دلاليَّة لها معنى.
- ✓ اختلفَ العلماءُ حولَ تمييز الوحدة الدِّلاليَّة وتعيين حُدودها؛ فرأى البعضُ أنَّها تُرادفُ (الكلمة)، ورأى آخرونَ أنَّها جزءٌ من الكلمة.
- ✓ يُساعد التَّحليلُ الدِّلاليُّ في التَّعرُّف على العلاقات بين المُفردات، أو ما يُعرَف بـ (العلاقات الدِّلاليَّة).

مُستويات التَّحليل اللُّغويِّ في حوسِّبة اللُّغة

• التَّحليل الدَّلاليّ:

علاقة التَّرادُف *Synonym*: وَمِنْ أُمَثِلَتِهَا فِي الْعَرَبِيَّةِ:

(رَجُل) ← (ذَكَر / بَالِغ / شُجَاع).

علاقة التَّضادِّ *Antonymy*: وَمِنْ أُمَثِلَتِهَا فِي الْعَرَبِيَّةِ:

(التَّضادُّ الحَقِيقِيّ) ← (الصِّدْق، الكَذِب / الذَّكَر، الأنثى).

(التَّضادُّ العَكْسِيّ) ← (بَاعَ، اشْتَرَى / زَوْج، زَوْجَة).

علاقة الاحتواء *Hyponym*: وَمِنْ أُمَثِلَتِهَا فِي الْعَرَبِيَّةِ:

(أُسْرَة) ← (أَب / أُم / ابْن / ابْنَة).

علاقة الانتماء *Hypernym*: وَمِنْ أُمَثِلَتِهَا فِي الْعَرَبِيَّةِ:

(أَب / أُم / ابْن / ابْنَة) ← (أُسْرَة).



شُكْرًا لَكُمْ

للتواصل ..

• د. المعتز بالله السعيد

moataz@cu.edu.eg

almoataz2006@skype.com

[almo3tazbellah@facebook.com](https://www.facebook.com/almo3tazbellah)

www.ask.fm/AlmoatazBAISaid

<http://scholar.cu.edu.eg/moataz>

<http://cairo.academia.edu/AlmoatazBAISaid>